Linear Regression

ISC Class 12

Previous Years Board Questions

with Answers

2005 to 2013

1. There are two series of index numbers: P for price index and S for stock of a commodity. The mean and the standard deviation of P are 100 and 8 and of S are 103 and 4 respectively. The correlation coefficient between the two series is 0.4. With these data, obtain the regression lines of P on S and S on P.

(ISC 2005)

2. Out of the two regression lines, find the line of regression of y on x:

$$3x + 12y = 9,\ 9x + 3y = 46$$

(ISC 2005)

3. The following table shows the sales and expenditure of a farm:

	Sales (₹ in Crores)	advertisement expenditure (₹ in Crores)
Mean	40	6
Standard deviation	10	1.5

coefficient of correlation = $r = 0.9$

Estimate the likely sells for a proposed advertisement expenditure of ₹ 10 crores.

(ISC 2006)

4. Given two regression lines $4x + 3y + 7 = 0$ and $3x + 4y + 8 = 0$, determine

(i) the regression line of y on x.

(ii) the regression line of x on y.

(iii) the coefficient of correlation. (ISC 2007)

5. The data for marks in Physics and History obtained by 10 students are given below:

Physics	15	12	8	8	7	7	7	6	5	3
History	10	25	17	11	13	17	20	13	9	15

using this data hold on:

(i) find the line of regression in which Physics is taken as a independent variable.

(ii) A candidate had scored 10 marks in Physics but was absent from the History test. Estimate his probable score for the latter test. (ISC 2008)

6.	If the two regression lines of a bivariate distribution are $4x - 5y + 33 = 0$ and $20x - 9y - 107 = 0$, (i) Calculate, $\bar{x}$ and $\bar{y}$, the arithmetic means of x and y respectively. (ii) Estimate the value of x when $y = 7$. (iii) Find the variance of y when $\sigma_x = 3$. (ISC 2009)
7.	If the regression equation of x on y is given by $mx - y + 10 = 0$ and the equation of y on x is given by $-2x + 5y = 14$, determine the value of $'m'$ if the coefficient of correlation between x & y is $\frac{1}{\sqrt{10}}$. (ISC 2010)
8.	Find the equations of the two lines of regression for the following observations: (3, 6), (4, 5), (5, 4), (6, 3), (7, 2) Find an estimate of y for $x = 2.5$. (ISC 2010)
9.	Two regression lines are represented by $2x + 3y - 10 = 0$ and $4x + y - 5 = 0$. Find the line of regression of y on x. (ISC 2011)
10.	The following observations are given: (1, 4), (2, 8), (3, 2), (4, 12), (5, 10), (6, 14), (7, 16), (8, 6), (9, 18) Estimate the value of y when the value of x is 10 add also estimate the value of x when the value of y is 5. (ISC 2011)
11.	From the equations of the two regression lines $4x + 3y + 7 = 0$ and $3x + 4y + 8 = 0$, find (i) the mean values of x and y (ii) regression coefficient (iii) coefficient of correlation (ISC 2012)
12.	The following results were obtained with respect to two variables x and y: $\sum x = 30 . \sum y = 42, \sum xy = 199, \sum x^2 = 184, \sum y^2 = 318, n = 6$ Find the following: (i) the regression coefficients (ii) the correlation coefficient between x and y. (iii) the regression equation of y on x. (iv) the likely value of y when $x = 10$. (ISC 2012)
13.	Two Regression lines are represented by $4x + 10y = 9$ and $6x + 3y = 4$. Find Regression line of y on x. (ISC 2013)
14.	The following table shows the mean and the standard deviation of the marks of Mathematics and Physics scored by the students of a school: | | Mathematics | Physics | | Mean | 84 | 81 | | standard deviation | 7 | 4 | The correlation coefficient between the given marks is 0.86. Estimate the likely marks in Physics if the marks in Mathematics are 92. (ISC 2013)

Answer

1.	Regression lines of P on S is P = 0.8 S +17.6 and regression line of S on P is S = 0.2 P + 83	(ISC 2005)
2.	$3x + 12y = 9$	(ISC 2005)
3.	₹ 64 crores	(ISC 2006)
4.	(i) $3x + 4y + 8 = 0$ (ii) $4x + 3y + 7 = 0$ (iii) $r = -\frac{3}{4}$	(ISC 2007)
5.	Take the marks obtained in Physics and History as x and y respectively. (i) $24\,y = 5x + 321$ (ii) Probable score in history = 15.5 marks	(ISC 2008)
6.	(i) $\bar{x} = 13,\ \bar{y} = 17$ (ii) When $y = 7,\ x = 8.5$ (iii) The variance of $y = 16$	(ISC 2009)
7.	$m = 4$	(ISC 2010)
8.	Regression line of y on x is $y = -x + 9$. When $x = 2.5,\ y = 6.5$ Regression line of x on y is $x = -y + 9$.	(ISC 2010)
9.	$2x + 3y - 10 = 0$	(ISC 2011)
10.	When $x = 10$, then $y = \frac{50}{3}$; when $y = 5$, then $x = \frac{10}{3}$.	(ISC 2011)
11.	(i) $\bar{x} = -\frac{4}{7},\ \bar{y} = -\frac{11}{7}$ (ii) $b_{xy} = -\frac{3}{4},\quad b_{yx} = -\frac{3}{4}$ (iii) $r = -\frac{3}{4}$	(ISC 2012)
12.	(i) $b_{yx} = -\frac{11}{34} = -0.323,\ b_{xy} = -\frac{11}{24} = -0.458,$ (ii) $r = -\sqrt{\frac{121}{816}} = -0.385$ (iii) $11x + 34y - 293 = 0$ (iv) When $x = 10,\ y = 5.382$	(ISC 2012)
13.	$4x + 10y = 9$	(ISC 2013)
14.	84.93	(ISC 2013)

Linear Regression

ISC Class 12

Previous Years Board Questions with Answers

2014 to 2022

1.	For the given lines of regression, $3x - 2y = 5$ and $x - 4y = 7$, find: (i) regression coefficients b_{yx} and b_{xy} (ii) coefficient of correlation $r(x, y)$. (ISC 2014)
2.	The coefficient of correlation between the values denoted by X and Y is 0.5. The mean of X is 3 and that of Y Is 5. Their standard deviation are 5 and 4 respectively. Find: (i) the two lines of regression (ii) the expected value of Y, when X Is given 14 (iii) the expected value of X , when Y Is given 9. (ISC 2014)
3.	The two lines of regression are $4x + 2y - 3 = 0$ and $3x + 6y + 5 = 0$. Find the correlation coefficient between x and y. (ISC 2015)
4.	Given that the observations are: (9, – 4), (10, – 3) , (11, – 1), (12,0), (13, 1), (14, 3), (15, 5), (16, 8). Find the two lines of regression and estimate the value of y when $x = 13.5$ (ISC 2015)
5.	The two lines of regression are $x + 2y - 5 = 0$ and $2x + 3y - 8 = 0$ and the variance of x is 12. Find the variance of y and the coefficient of correlation. (ISC 2016)
6.	By using the data $\bar{x} = 25, \bar{y} = 30, b_{yx} = 1.6, b_{xy} = 0.4$, find: (a) The regression equation of y on x. (b) What is the most likely value of y when $x = 60$? (c) What is the coefficient of correlation between x and y? (ISC 2017)

7. Find the line of best fit for the following data, treating x as dependent variable (regression equation of x on y):

x	14	12	13	14	16	10	13	12
y	14	23	17	24	18	25	23	24

hence, estimate the value of x when $y = 16$. (ISC 2017)

8. Find the line of regression of y on x from the following data:

x	1	2	3	4	5
y	7	6	5	4	3

Hence, estimate the value of y when $x = 6$. (ISC 2018)

9. Find the coefficient of correlation from the following regression lines:
$x - 2y + 3 = 0, 4x - 5y + 1 = 0$ (ISC 2018)

10. From the given data

	x	y
Mean	6	8
standard deviation	4	6

and coefficient correlation $\frac{2}{3}$. Find:

(i) regression coefficients b_{yx} and b_{xy}.

(ii) regression line of x on y.

(iii) most likely values of x when $y = 14$. (ISC 2018)

11. If $\bar{x} = 18, \bar{y} = 100, \sigma_x = 14,\ \sigma_y = 20$ and correlation coefficient $r_{xy} = 0.8$, find the regression equation of y on x. (ISC 2019)

12. The following results were obtained with respect to two variables x and y:
$\sum x = 15,\ \sum y = 25, \sum xy = 83,\ \sum x^2 = 55, \sum y^2 = 135$ and $n = 5$

(i) Find the regression coefficient b_{xy}

(ii) Find the regression equation of x on y. (ISC 2019)

13. Find the equation of the regression line of y on x, if the observations (x, y) are as follows:

(1,4), (2,8), (3,2), (4,12), (5,10), (6,14), (7,16), (8,6), (9,18).

Also, find the estimated value of y when $x = 14$. (ISC 2019)

14. For the lines of regression $4x - 2y = 4$ and $2x - 3y + 6 = 0$, find the mean of $'x'$, mean of $'y'$. (ISC 2020)

15.	The correlation coefficient between x and y Is 0.6. If the variance of x is 225, the variance of y is 400, mean of x is 10 and mean of y is 20, find: (i) The equations of 2 regression lines (ii) The expected value of y when $x = 2$. (ISC 2020)
16.	Find the regression coefficients b_{yx} and b_{xy} and correlation coefficient $'r'$ for the following data: (2,8), (6, 8), (4,5), (7,6), (5,2) (ISC 2020)
17.	Choose the correct option: (ISC SEM-2 2022) If the two regression coefficients b_{xy} and b_{yx} are -0.8 and -0.2 respectively, then the correlation coefficient (ρ) will be, (a) 0.16 (b) – 0.16 (c) 0.4 (d) – 0.4
18.	Choose the correct option: (ISC SEM-2 2022) The line of regression of y on x is, $4x - 5y + 33 = 0$ and the line of regression of x on y is, $20x - 9y - 107 = 0$, then the value of x when $y = 7$ is, (a) 8.5 (b) – 8.5 (c) 0.5 (d) – 0.5
19.	The mean and standard deviation of the two variables x and y are given as $\bar{x} = 6,\ \bar{y} = 8,\ \sigma_x = 4,\ \sigma_y = 12$. The correlation coefficient is given as $r = \frac{2}{3}$. Find the regression line of x on y. (ISC SEM-2 2022)

Answer

1.	(i) $b_{yx} = \frac{1}{4},\ b_{xy} = \frac{2}{3}$ (ii) $r = \frac{1}{\sqrt{6}} = 0.408$	(ISC 2014)
2.	(i) regression line of y on x is $2x - 5y + 19 = 0$ and regression line of x on y is $8x - 5y + 1 = 0$ (ii) When $x = 14,\ y = 9.4$ (iii) When $y = 9,\ x = 5.5$	(ISC 2014)
3.	$r = -\frac{1}{2}$	(ISC 2015)

4.	Regression line of y on x is $y = 1.63\,x - 19.25$ and regression line of x on y is $x = 0.596\,y + 11.83$. When $x = 13.5,\ y = 2.76.$	(ISC 2015)
5.	The variance of $y = 4$, coefficient of correlation = $-\frac{\sqrt{3}}{2} = -0.866$	(ISC 2016)
6.	(a) $8x - 5y - 50 = 0$ (b) $y = 86$ (c) 0.8	(ISC 2017)
7.	The line of regression of x on y is $58\,x + 15\,y = 1069.$ When $y = 16, x = 14.29$	(ISC 2017)
8.	$x + y = 8$, when $x = 6, y = 2$	(ISC 2018)
9.	0.79	(ISC 2018)
10.	(i) $b_{yx} = 1, \quad b_{xy} = \frac{4}{9}$ (ii) $9x = 4y + 22$ (iii) 8.667	(ISC 2018)
11.	$8\,x - 7y + 556 = 0$	(ISC 2019)
12.	(i) $b_{xy} = \frac{4}{5}$ (ii) $5x - 4y + 5 = 0$	(ISC 2019)
13.	$4x - 3y + 10 = 0$, when $x = 14,\ y = 22$	(ISC 2019)
14.	$\bar{x} = 3,\ \bar{y} = 4$	(ISC 2020)
15.	The regression equation of y on x is $4x - 5y + 60 = 0$ The regression equation of x on y is $20\,x - 9y - 20 = 0.$ When $x = 2, \quad y = 13.6$	(ISC 2020)
16.	$b_{yx} = -0.216, \quad b_{xy} = -0.129, \quad r = -0.1669$	(ISC 2020)
17.	Option (d) -0.4	(ISC 2022)
18.	Option (a) 8.5	(ISC 2022)
19.	$9x - 2y = 38$	(ISC 2022)

Solutions of

Linear Regression

ISC Previous Years Board Questions

2005 to 2013

1. There are two series of index numbers: P for price index and S for stock of a commodity. The mean and the standard deviation of P are 100 and 8 and of S are 103 and 4 respectively. The correlation coefficient between the two series is 0.4. With these data, obtain the regression lines of P on S and S on P. (ISC 2005)

Solution: Given

Commodity	Price P (x)	Stock S (y)
Mean	$\bar{x} = 100$	$\bar{y} = 103$
Standard deviation	$\sigma_x = 8$	$\sigma_y = 4$
Correlation coefficient	$r = 0.4$	

Regression coefficient of S on P is $b_{yx} = r\,\frac{\sigma_y}{\sigma_x} = 0.4 \times \frac{4}{8} = 0.2$

Regression coefficient of P on S is $b_{xy} = r\,\frac{\sigma_x}{\sigma_y} = 0.4 \times \frac{8}{4} = 0.8$

Hence the regression line of P on S is (i.e., regression equation of x on y)

$$x - \bar{x} = b_{xy}(y - \bar{y})$$
$$\Rightarrow x - 100 = 0.8\,(y - 103)$$
$$\Rightarrow x - 100 = 0.8\,y - 82.4$$
$$\Rightarrow x = 0.8\,y + 17.6$$

The regression line of S on P is (i.e., regression equation of y on x)

$$y - \bar{y} = b_{yx}(x - \bar{x})$$
$$\Rightarrow y - 103 = 0.2\,(x - 100)$$
$$\Rightarrow y = 0.2\,x - 20 + 103$$
$$\Rightarrow y = 0.2\,x + 83$$

2. Out of the two regression lines, find the line of regression of y on x:

$$3x + 12y = 9,\ 9x + 3y = 46$$ (ISC 2005)

Solution: Let equation $3x + 12y = 9$ is the regression equation of y on x and equation $9x + 3y = 46$ is the regression equation of x on y.

$$3x + 12y = 9$$
$$\Rightarrow 12\,y = -3x + 9$$
$$\Rightarrow y = -\frac{3}{12}\,x + \frac{9}{12}$$
$$\Rightarrow y = -\frac{1}{4}\,x + \frac{3}{4}$$
$$\therefore b_{yx} = -\frac{1}{4}$$

$$9x + 3y = 46$$
$$\Rightarrow 9x = -3y + 46$$
$$\Rightarrow x = -\frac{3}{9}\,y + \frac{46}{9}$$
$$\Rightarrow x = -\frac{1}{3}\,y + \frac{46}{9}$$
$$\therefore b_{xy} = -\frac{1}{3}$$

$\therefore r^2 = -\frac{1}{4} \times -\frac{1}{3} = \frac{1}{12}$ $\left[0 < \frac{1}{12} < 1\right]$

We know $0 \le r^2 \le 1$

So our assumption is correct.

Hence the equation $3x + 12y = 9$ is the regression equation of y on x.

3. The following table shows the sales and expenditure of a farm:

	Sales (₹ in Crores)	advertisement expenditure (₹ in Crores)
Mean	40	6
Standard deviation	10	1·5

coefficient of correlation = $r = 0{\cdot}9$

Estimate the likely sells for a proposed advertisement expenditure of ₹ 10 crores. (ISC 2006)

Solution:

	Sales (₹ in Crores) (x)	advertisement expenditure (₹ in Crores) (y)
Mean	$\bar{x} = 40$	$\bar{y} = 6$
Standard deviation	$\sigma_x = 10$	$\sigma_y = 1{\cdot}5$

And $r = 0{\cdot}9$

Regression coefficient of x on $y = b_{xy} = r \cdot \frac{\sigma_x}{\sigma_y} = 0{\cdot}9 \cdot \frac{10}{1{\cdot}5} = 6$

$\therefore$ regression equation of x on y is $x - \bar{x} = b_{xy}(y - \bar{y})$

$\Rightarrow x - 40 = 6(y - 6)$

$\Rightarrow x = 6y - 36 + 40$

$\Rightarrow x = 6y + 4$

$\therefore$ when $y = 10$ crores, sales $x = 6 \times 10 + 4 = 64$ crores.

4. Given two regression lines $4x + 3y + 7 = 0$ and $3x + 4y + 8 = 0$, determine

(i) the regression line of y on x.

(ii) the regression line of x on y.

(iii) the coefficient of correlation. (ISC 2007)

Solution:

Let equation $4x + 3y + 7 = 0$ is the regression equation of x on y and equation $3x + 4y + 8 = 0$ is the regression equation of y on x.

$4x + 3y + 7 = 0$

$\Rightarrow 4x = -3y - 7$

$\Rightarrow x = -\frac{3}{4}y - \frac{7}{4}$

$\therefore b_{xy} = -\frac{3}{4}$

$3x + 4y + 8 = 0$

$\Rightarrow 4y = -3x - 8$

$\Rightarrow y = -\frac{3}{4}x - \frac{8}{4}$

$\therefore b_{yx} = -\frac{3}{4}$

$\therefore r^2 = b_{xy} \times b_{yx} = -\frac{3}{4} \times -\frac{3}{4} = \frac{9}{16}$ $\left[0 < \frac{9}{16} < 1\right]$

We know $0 \leq r^2 \leq 1$

$\therefore$ our assumption is correct.

(i) The equation $3x + 4y + 8 = 0$ is the regression equation of y on x.

(ii) The equation $4x + 3y + 7 = 0$ is the regression equation of x on y

(iii) $r^2 = \frac{9}{16} \Rightarrow r = -\sqrt{\frac{9}{16}} = -\frac{3}{4}$

$\because$ both b_{yx} and b_{xy} are negative, the correlation coefficient r is also negative.

5. The data for marks in Physics and History obtained by 10 students are given below:

Physics	15	12	8	8	7	7	7	6	5	3
History	10	25	17	11	13	17	20	13	9	15

Using this data:

(i) find the line of regression in which Physics is taken as a independent variable.

(ii) A candidate had scored 10 marks in Physics but was absent from the history test. Estimate his probable score for the latter test. (ISC 2008)

Solution: Let the marks obtained in Physics and History be x and y respectively.
Let assumed mean of marks of Physics be $= 8$ and assumed mean of marks of History be $= 15$.

x	y	$u = x - 8$	$v = y - 15$	$uv = (x-8)(y-15)$	$u^2 = (x-8)^2$
15	10	7	-5	-35	49
12	25	4	10	40	16
8	17	0	2	0	0
8	11	0	-4	0	0
7	13	-1	-2	2	1
7	17	-1	2	-2	1
7	20	-1	5	-5	1
6	13	-2	-2	4	4
5	9	-3	-6	18	9
3	15	-5	0	0	25
$\sum x = 78$	$\sum y = 150$	$\sum u = -2$	$\sum v = 0$	$\sum uv = 22$	$\sum u^2 = 106$

$\therefore \bar{x} = \frac{\sum x}{n} = \frac{78}{10} = 7.8$ and $\bar{y} = \frac{\sum y}{n} = \frac{150}{10} = 15$

(i) Since Physics is taken as a independent variable, we have to find out the regression line of y on x.

$$b_{yx} = \frac{\sum uv - \frac{1}{n}\sum u \sum v}{\sum u^2 - \frac{1}{n}(\sum u)^2} = \frac{22 - \frac{1}{10}(-2)(0)}{106 - \frac{1}{10}\times(-2)^2} = \frac{22}{105.6} = \frac{5}{24}$$

The regression line of y on x is

$$y - \bar{y} = b_{yx}(x - \bar{x})$$
$$\Rightarrow y - 15 = \frac{5}{24}(x - 7.8)$$
$$\Rightarrow 24y - 360 = 5x - 39$$
$$\Rightarrow 24y = 5x + 321$$ which is the equation of regression line taking Physics as independent variable.

(ii) When marks of Physics is 10 i.e., $x = 10$,

$$y = \frac{5}{24}x + \frac{321}{24} = \frac{5}{24}\times 10 + \frac{321}{24} = 15.5 \text{ (approximately)}$$

$\therefore$ The required probable score of History = 15.5 marks.

6. If the two regression lines of a bivariate distribution are $4x - 5y + 33 = 0$ and $20x - 9y - 107 = 0$,

(i) Calculate, $\bar{x}$ and $\bar{y}$, the arithmetic means of x and y respectively.

(ii) Estimate the value of x when $y = 7$.

(iii) Find the variance of y when $\sigma_x = 3$. (ISC 2009)

Solution:

(i) $4x - 5y + 33 = 0$ ---------- (1)

$20x - 9y - 107 = 0$ --------- (2)

Multiplying equation (1) by 5,

$$\begin{array}{r} 20x - 25y + 165 = 0 \\ 20x - 9y - 107 = 0 \\ - \quad + \quad + \\ \hline -16y + 272 = 0 \end{array} \quad \Rightarrow y = 17 \text{ i.e., } \bar{y} = 17$$

Putting $y = 17$ in (1), $x = 13$, i.e., $\bar{x} = 13$

$\therefore \quad \bar{x} = 13, \quad \bar{y} = 17$

(ii) Let equation (1) is the regression equation of y on x and equation (2) is the regression equation of x on y.

$4x - 5y + 33 = 0$

$\Rightarrow \quad 5y = 4x + 33$

$\Rightarrow \quad y = \frac{4}{5}x + \frac{33}{5}$

$\therefore \quad b_{yx} = \frac{4}{5}$

$20x - 9y - 107 = 0$

$\Rightarrow \quad 20x = 9y + 107$

$\Rightarrow \quad x = \frac{9}{20}y + \frac{107}{20}$

$\therefore \quad b_{xy} = \frac{9}{20}$

$\therefore \quad r^2 = b_{yx} \times b_{xy} = \frac{4}{5} \times \frac{9}{20} = \frac{9}{25}$ $\quad \left[0 < \frac{9}{25} < 1\right]$ (Assumption correct)

$\therefore r = \sqrt{\frac{9}{25}} = \frac{3}{5}$ (r is positive as both b_{yx}, b_{xy} are positive)

$\therefore$ the regression equation of x on y is $20x - 9y - 107 = 0$

$\Rightarrow \quad 20x = 9y + 107$

$\Rightarrow \quad x = \frac{9}{20}y + \frac{107}{20}$

When $y = 7, \quad x = \frac{9}{20} \times 7 + \frac{107}{20} = \frac{170}{20} = 8{\cdot}5$

(iii) $b_{yx} = r \cdot \frac{\sigma_y}{\sigma_x} \quad \Rightarrow \frac{4}{5} = \frac{3}{5} \times \frac{\sigma_y}{3} \quad \Rightarrow \quad \sigma_y = 4$

$\therefore$ the variance of $y = \sigma_y^2 = 16$

---

7. If the regression equation of x on y is given by $mx - y + 10 = 0$ and the equation of y on x is given by $-2x + 5y = 14$, determine the value of 'm' if the coefficient of correlation between x & y is $\frac{1}{\sqrt{10}}$. (ISC 2010)

Solution:

The regression equation of x on y is given by $mx - y + 10 = 0$

$\Rightarrow \quad mx = y - 10$

$\Rightarrow \quad x = \frac{1}{m}y - \frac{10}{m}$

$\therefore \quad b_{xy} = \frac{1}{m}$

The regression equation of y on x is given by $-2x + 5y = 14$

$\Rightarrow \quad 5y = 2x + 14$

$\Rightarrow \quad y = \frac{2}{5}x + \frac{14}{5}$

$\therefore \quad b_{yx} = \frac{2}{5}$

Given $r = \frac{1}{\sqrt{10}}$. We know $r^2 = b_{yx} \times b_{xy}$

$$\therefore \quad \frac{1}{10} = \frac{1}{m} \times \frac{2}{5}$$

$$\therefore m = 4$$

8. Find the equations of the two lines of regression for the following observations:

(3, 6), (4, 5), (5, 4), (6, 3), (7, 2)

Find an estimate of y for $x = 2.5$. (ISC 2010)

Solution:

x	y	x^2	y^2	xy
3	6	9	36	18
4	5	16	25	20
5	4	25	16	20
6	3	36	9	18
7	2	49	4	14
$\sum x = 25$	$\sum y = 20$	$\sum x^2 = 135$	$\sum y^2 = 90$	$\sum xy = 90$

$n = 5, \quad \bar{x} = \frac{\sum x}{n} = \frac{25}{5} = 5, \qquad \bar{y} = \frac{\sum y}{n} = \frac{20}{5} = 4$

Regression coefficient of y on $x = b_{yx} = \frac{\sum xy - \frac{\sum x \cdot \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} = \frac{90 - \frac{25 \times 20}{5}}{135 - \frac{25^2}{5}} = \frac{-10}{10}$

$= -1$

Regression coefficient of x on $y = b_{xy} = \frac{\sum xy - \frac{\sum x \cdot \sum y}{n}}{\sum y^2 - \frac{(\sum y)^2}{n}} = \frac{90 - \frac{25 \times 20}{5}}{90 - \frac{20^2}{5}} = \frac{-10}{10}$

$= -1$

The regression line of y on x is $y - \bar{y} = b_{yx}(x - \bar{x})$

$\Rightarrow \quad y - 4 = -1(x - 5)$

$\Rightarrow \quad x + y = 9$

The regression line of of x on y is given by $(x - \bar{x}) = b_{xy}(y - \bar{y})$

$\Rightarrow \quad x - 5 = -1(y - 4)$

$\Rightarrow \quad x + y - 9 = 0$

We have to estimate y when $x = 2.5$.

The regression line of y on x is $x + y = 9$

$\Rightarrow \quad y = -x + 9$

$\Rightarrow \quad y = -2.5 + 9 = 6.5$

9. Two regression lines are represented by $2x + 3y - 10 = 0$ and $4x + y - 5 = 0$. Find the line of regression of y on x. (ISC 2011)

Solution: Let $2x + 3y - 10 = 0$ is the regression line of y on x and $4x + y - 5 = 0$ is the regression line of x on y.

$2x + 3y - 10 = 0$

$\Rightarrow \quad 3y = -2x + 10$

$\Rightarrow \quad y = -\frac{2}{3}x + \frac{10}{3}$

$\therefore \quad b_{yx} = -\frac{2}{3}$

$4x + y - 5 = 0$

$\Rightarrow \quad 4x = -y + 5$

$\Rightarrow \quad x = -\frac{1}{4}y + \frac{5}{4}$

$\therefore \quad b_{xy} = -\frac{1}{4}$

$$r^2 = b_{yx} \times b_{xy} = \left(-\frac{2}{3}\right)\left(-\frac{1}{4}\right) = \frac{1}{6}, \qquad 0 < \frac{1}{6} < 1$$

We know $0 \leq r^2 \leq 1$, so our assumption is correct.

Hence $2x + 3y - 10 = 0$ is the regression line of y on x.

10. The following observations are given:

(1, 4), (2, 8), (3, 2), (4, 12), (5, 10), (6, 14), (7, 16), (8, 6), (9, 18)

Estimate the value of y when the value of x is 10 add also estimate the value of x when the value of y is 5. (ISC 2011)

Solution:

$$\bar{x} = \frac{\sum x}{n} = \frac{45}{9} = 5 \text{ and } \bar{y} = \frac{\sum y}{n} = \frac{90}{9} = 10$$

x	y	$x - \bar{x} = x - 5$	$y - \bar{y} = y - 10$	$(x-x)^2 = (x-5)^2$	$(y-\bar{y})^2 = (y-10)^2$	$(x-\bar{x})(y-\bar{y})$
1	4	-4	-6	16	36	24
2	8	-3	-2	9	4	6
3	2	-2	-8	4	64	16
4	12	-1	2	1	4	-2
5	10	0	0	0	0	0
6	14	1	4	1	16	4
7	16	2	6	4	36	12
8	6	3	-4	9	16	-12
9	18	4	8	16	64	32
$\sum x = 45$	$\sum y = 90$			$\sum(x-\bar{x})^2 = 60$	$\sum(y-\bar{y})^2 = 240$	$\sum(x-\bar{x})(y-\bar{y}) = 80$

$$b_{yx} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2} = \frac{80}{60} = \frac{4}{3}$$

$$b_{xy} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(y-\bar{y})^2} = \frac{80}{240} = \frac{1}{3}$$

The regression equation of y on x is

$$y - \bar{y} = b_{yx}(x - \bar{x})$$

$$\Rightarrow \quad y - 10 = \frac{4}{3}(x - 5)$$

$$\Rightarrow \quad y = \frac{4}{3}x - \frac{20}{3} + 10$$

$$\Rightarrow \quad y = \frac{4}{3}x + \frac{10}{3}$$

When $x = 10$, $y = \frac{4}{3} \times 10 + \frac{10}{3} = \frac{50}{3}$

The regression equation of x on y is

$$x - \bar{x} = b_{xy}(y - \bar{y})$$

$$\Rightarrow \quad x - 5 = \frac{1}{3}(y - 10)$$

$$\Rightarrow \quad x = \frac{1}{3}y + 5 - \frac{10}{3}$$

$$\Rightarrow \quad x = \frac{1}{3}y + \frac{5}{3}$$

When $y = 5$, $x = \frac{5}{3} + \frac{5}{3} = \frac{10}{3}$

11. From the equations of the two regression lines $4x + 3y + 7 = 0$ and $3x + 4y + 8 = 0$, find:

(i) the mean values of x and y

(ii) regression coefficient.

(iii) Coefficient of correlation (ISC 2012)

Solution: Given equations of two regression lines

$4x + 3y + 7 = 0$ ---------- (1)

$3x + 4y + 8 = 0$ ---------- (2)

From (1) $16x + 12y + 28 = 0$

From (2) $9x + 12y + 24 = 0$

$7x + 4 = 0 \quad \Rightarrow x = -\frac{4}{7}$

Putting $x = -\frac{4}{7}$ in (1), $4 \times -\frac{4}{7} + 3y + 7 = 0 \quad \Rightarrow 3y = -7 + \frac{16}{7}$

$\Rightarrow 3y = -\frac{33}{7} \quad \Rightarrow y = -\frac{11}{7}$

(i) $\bar{x} = -\frac{4}{7}$, $\bar{y} = -\frac{11}{7}$

(ii) Let equation $4x + 3y + 7 = 0$ be the regression line of x on y and equation $3x + 4y + 8 = 0$ be the regression line of y on x.

$4x + 3y + 7 = 0$

$\Rightarrow 4x = -3y - 7$

$\Rightarrow x = -\frac{3}{4}y - \frac{7}{4}$

$\therefore b_{xy} = -\frac{3}{4}$

$3x + 4y + 8 = 0$

$\Rightarrow 4y = -3x - 8$

$\Rightarrow y = -\frac{3}{4}x - 2$

$\therefore b_{yx} = -\frac{3}{4}$

$r^2 = b_{xy} \times b_{yx} = \left(-\frac{3}{4}\right)\left(-\frac{3}{4}\right) = \frac{9}{16} \qquad \left(0 < \frac{9}{16} < 1\right)$

We know $0 \leq r^2 \leq 1$. So our assumption is correct.

$\therefore b_{xy} = -\frac{3}{4}$ and $b_{yx} = -\frac{3}{4}$

(iii) $r^2 = \frac{9}{16} \quad \Rightarrow r = -\frac{3}{4}$ (as both b_{xy}, b_{yx} are negative, coefficient of correlation r is also negative.)

12. The following results were obtained with respect to two variables x and y:

$\sum x = 30 . \sum y = 42, \sum xy = 199, \sum x^2 = 184, \sum y^2 = 318, n = 6$

Find the following:

(i) the regression coefficients

(ii) the correlation coefficient between x and y.

(iii) the regression equation of y on x.

(iv) the likely value of y when $x = 10$.

(ISC 2012)

Solution: We have $\sum x = 30 . \sum y = 42, \sum xy = 199, \sum x^2 = 184, \sum y^2 = 318, n = 6$

(i)

Regression coefficient of y on $x = b_{yx} = \frac{\sum xy - \frac{\sum x . \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} = \frac{199 - \frac{30 \times 42}{6}}{184 - \frac{30^2}{6}} = -\frac{11}{34}$

Regression coefficient of x on $y = b_{xy} = \frac{\sum xy - \frac{\sum x . \sum y}{n}}{\sum y^2 - \frac{(y)^2}{n}} = \frac{199 - \frac{30 \times 42}{6}}{318 - \frac{42^2}{6}} = -\frac{11}{24}$

(ii) Since b_{yx} and b_{xy} both are negative, the correlation coefficient is also negative.

$$r = -\sqrt{b_{yx} \times b_{xy}} = -\sqrt{\left(-\frac{11}{34}\right) \times \left(-\frac{11}{24}\right)} = -\sqrt{\frac{121}{816}} = -0{\cdot}385$$

(iii) $\bar{x} = \frac{\sum x}{n} = \frac{30}{6} = 5$ and $\bar{y} = \frac{\sum y}{n} = \frac{42}{6} = 7$

∴The regression line of y on x is $y - \bar{y} = b_{yx}(x - \bar{x})$

$$\Rightarrow \quad y - 7 = -\frac{11}{34}(x - 5)$$

$$\Rightarrow \quad 34y - 238 = -11x + 55$$

$$\Rightarrow \quad 11x + 34y - 293 = 0$$

(iv) When $x = 10$, $110 + 34y - 293 = 0 \Rightarrow y = \frac{183}{34}$

Or $y = 5{\cdot}382$

13. Two Regression lines are represented by $4x + 10y = 9$ and $6x + 3y = 4$. Find Regression line of y on x. (ISC 2013)

Solution: Let equation $4x + 10y = 9$ is the regression equation of y on x and equation $6x + 3y = 4$ is the regression equation of x on y.

$4x + 10y = 9$

$\Rightarrow \quad 10y = -4x + 9$

$\Rightarrow \quad y = -\frac{4}{10}x + \frac{9}{10}$

$\therefore \quad b_{yx} = -\frac{4}{10} = -\frac{2}{5}$

$6x + 3y = 4$

$\Rightarrow \quad 6x = -3y + 4$

$\Rightarrow \quad x = -\frac{3}{6}y + \frac{4}{6}$

$\therefore \quad b_{xy} = -\frac{3}{6} = -\frac{1}{2}$

$$\therefore \quad r^2 = b_{yx} \times b_{xy} = \left(-\frac{2}{5}\right)\left(-\frac{1}{2}\right) = \frac{1}{5} \qquad \left[0 < \frac{1}{5} < 1\right]$$

We know, $0 \leq r^2 \leq 1$. So our assumption is correct.

Hence, the equation $4x + 10y = 9$ is the regression equation of y on x.

14. The following table shows the mean and the standard deviation of the marks of Mathematics and Physics scored by the students of a school:

	Mathematics	Physics
Mean	84	81
standard deviation	7	4

The correlation coefficient between the given marks is 0.86. Estimate the likely marks in Physics if the marks in Mathematics are 92. (ISC 2013)

Solution: Given

	Mathematics (x)	Physics (y)
Mean	$\bar{x} = 84$	$\bar{y} = 81$
standard deviation	$\sigma_x = 7$	$\sigma_y = 4$

And $r = 0.86$

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = 0.86 \times \frac{4}{7} = \frac{86}{175}$$

The regression line of y on x is

$$y - \bar{y} = b_{yx}(x - \bar{x})$$

$$\Rightarrow \quad y - 81 = \frac{86}{175}(x - 84)$$

$$\Rightarrow \quad 175y - 14175 = 86x - 7224$$

$$\Rightarrow \quad 175y = 86x + 6951$$

When $x = 92$, $175y = 86 \times 92 + 6951 = 14863$

$$\therefore \quad y = \frac{14863}{175} = 84.93$$

So when marks in Mathematics are 92, likely marks in Physics $= 84.93$

Solutions of

Linear Regression

ISC Previous Years Board Questions

2014 to 2022

1. For the given lines of regression, $3x - 2y = 5$ and $x - 4y = 7$, find:

(i) regression coefficients b_{yx} and b_{xy}

(ii) coefficient of correlation $r(x, y)$. (ISC 2014)

Solution:

(i) let $3x - 2y = 5$ be the regression equation of x on y and $x - 4y = 7$ is the regression equation of y on x.

$3x - 2y = 5$

$\Rightarrow 3x = 2y + 5$

$\Rightarrow x = \frac{2}{3}y + \frac{5}{3}$

$\therefore b_{xy} = \frac{2}{3}$

$x - 4y = 7$

$\Rightarrow 4y = x - 7$

$\Rightarrow y = \frac{1}{4}x - \frac{7}{4}$

$\therefore b_{yx} = \frac{1}{4}$

$\therefore r^2 = b_{xy} \times b_{yx} = \frac{2}{3} \times \frac{1}{4} = \frac{1}{6}$ $\left[0 < \frac{1}{6} < 1\right]$

We know $0 \leq r^2 \leq 1$. So our assumption is correct.

$\therefore b_{xy} = \frac{2}{3}$ and $b_{yx} = \frac{1}{4}$

(ii) $r^2 = \frac{1}{6} \Rightarrow r = \frac{1}{\sqrt{6}}$ (r is positive as both b_{xy}, b_{yx} are positive)

Coefficient of correlation $r = \frac{1}{\sqrt{6}} = 0{\cdot}408$

2. The coefficient of correlation between the values denoted by X and Y is 0.5. The mean of X is 3 and that of Y Is 5. Their standard deviation are 5 and 4 respectively. Find:

(i) the two lines of regression

(ii) the expected value of Y, when X Is given 14

(iii) the expected value of X, when Y Is given 9. (ISC 2014)

Solution:

Given

	(X)	(Y)
Mean	$\bar{x} = 3$	$\bar{y} = 5$
standard deviation	$\sigma_x = 5$	$\sigma_y = 4$

And $r = 0.5$

(i) $b_{xy} = r\,\frac{\sigma_x}{\sigma_y} = 0.5 \times \frac{5}{4} = \frac{5}{8} = 0.625$

$b_{yx} = r\,\frac{\sigma_y}{\sigma_x} = 0.5 \times \frac{4}{5} = \frac{2}{5} = 0.4$

$\therefore$ the regression equation of x on y is,

$x - \bar{x} = b_{xy}(y - \bar{y})$

$\Rightarrow x - 3 = \frac{5}{8}(y - 5)$

$\Rightarrow 8x = 5y - 1$

The regression equation of y on x is

$y - \bar{y} = b_{yx}(x - \bar{x})$

$\Rightarrow \quad y - 5 = \frac{2}{5}(x - 3)$

$\Rightarrow \quad 5y = 2x + 19$

$\Rightarrow \quad 5y = 2x + 19$

(ii) When $X = 14$, $5Y = 2 \times 14 + 19 \quad \Rightarrow Y = 9.4$

(iii) $x = 0.625y - 0.125$

When $Y = 9$, $8X = 5 \times 9 - 1 \quad \Rightarrow \quad X = 5.5$

3. The two lines of regression are $4x + 2y - 3 = 0$ and $3x + 6y + 5 = 0$. Find the correlation coefficient between x and y. (ISC 2015)

Solution: Let equation $4x + 2y - 3 = 0$ is the regression equation of x on y and equation $3x + 6y + 5 = 0$ is the regression equation of y on x.

$4x + 2y - 3 = 0$	$3x + 6y + 5 = 0$
$\Rightarrow \quad 4x = -2y + 3$	$\Rightarrow \quad 6y = -3x - 5$
$\Rightarrow \quad x = -\frac{2}{4}y + \frac{3}{4}$	$\Rightarrow \quad y = -\frac{3}{6}x - \frac{5}{6}$
$\therefore \quad b_{xy} = -\frac{2}{4} = -\frac{1}{2}$	$b_{yx} = -\frac{3}{6} = -\frac{1}{2}$

$\therefore \quad r^2 = b_{yx} \times b_{xy} = \left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right) = \frac{1}{4} \quad [0 < \frac{1}{4} < 1]$

We know that $0 \leq r^2 \leq 1$. So our assumption is correct.

$\therefore \quad r = -\sqrt{\frac{1}{4}} = -\frac{1}{2}$ ($\because$ both b_{yx} and b_{xy} are negative, correlation coefficient r is also negative.)

4. Given that the observations are:

(9, – 4), (10, – 3) , (11, – 1), (12,0), (13, 1), (14, 3), (15, 5), (16, 8).

Find the two lines of regression and estimate the value of y when $x = 13.5$ (ISC 2015)

Solution:

x	y	x^2	y^2	xy
9	-4	81	16	-36
10	-3	100	9	-30
11	-1	121	1	-11
12	0	144	0	0
13	1	169	1	13
14	3	196	9	42
15	5	225	25	75
16	8	256	64	128
$\sum x = 100$	$\sum y = 9$	$\sum x^2 = 1292$	$\sum y^2 = 125$	$\sum xy = 181$

$n = 8, \quad \bar{x} = \frac{\sum x}{n} = \frac{100}{8} = 12{\cdot}5, \qquad \bar{y} = \frac{\sum y}{n} = \frac{9}{8} = 1{\cdot}125$

Regression coefficient of y on $x = b_{yx} = \frac{\sum xy - \frac{\sum x \cdot \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} = \frac{181 - \frac{100 \times 9}{8}}{1292 - \frac{100^2}{8}} = \frac{68{\cdot}5}{42}$

$= 1{\cdot}63$

Regression coefficient of x on $y = b_{xy} = \frac{\sum xy - \frac{\sum x \cdot \sum y}{n}}{\sum y^2 - \frac{(\sum y)^2}{n}} = \frac{181 - \frac{100 \times 9}{8}}{125 - \frac{9^2}{8}} = \frac{68{\cdot}5}{114.88}$

$= 0{\cdot}596$

The regression line of y on x is $\quad y - \bar{y} = b_{yx}\,(x - \bar{x})$

$\Rightarrow \quad y - 1{\cdot}125 = 1{\cdot}63\,(x - 12{\cdot}5)$

$\Rightarrow \quad y = 1{\cdot}63\,x - 19{\cdot}25$ ---------- (1)

The line of regression of x on y is given by $\quad (x - \bar{x}) = b_{xy}\,(y - \bar{y})$

$\Rightarrow \quad (x - 12{\cdot}5) = 0{\cdot}596\,(y - 1{\cdot}125)$

$\Rightarrow \quad x = 0{\cdot}596\,y + 11{\cdot}83$

When $x = 13{\cdot}5$, from equation (1), $\quad y = 1{\cdot}63 \times 13{\cdot}5 - 19{\cdot}25 = 2{\cdot}76$

5. The two lines of regression are $x + 2y - 5 = 0$ and $2x + 3y - 8 = 0$ and the variance of x is 12. Find the variance of y and the coefficient of correlation. (ISC 2016)

Solution: Let $x + 2y - 5 = 0$ is the regression line of y on x and $2x + 3y - 8 = 0$ is the regression line of x on y.

$x + 2y - 5 = 0$

$\Rightarrow 2y = -x + 5$

$\Rightarrow y = -\frac{1}{2}x + \frac{5}{2}$

$\therefore \; b_{yx} = -\frac{1}{2}$

$2x + 3y - 8 = 0$

$\Rightarrow \; 2x = -3y + 8$

$\Rightarrow x = -\frac{3}{2}y + 4$

$\therefore \; b_{xy} = -\frac{3}{2}$

$\therefore \; r^2 = b_{yx} \times b_{xy} = \left(-\frac{1}{2}\right) \times \left(-\frac{3}{2}\right) = \frac{3}{4} \qquad \left[0 < \frac{3}{4} < 1\right]$

We know $0 \le r^2 \le 1$. So our assumption is correct.

$\therefore \; r = \pm\frac{\sqrt{3}}{2}$ but b_{yx}, b_{xy} are both negative. So r is also negative.

The correlation coefficient $r = -\frac{\sqrt{3}}{2}$

Given ${\sigma_x}^2 = 12 \quad \Rightarrow \quad \sigma_x = \sqrt{12} = 2\sqrt{3}$

As $\quad b_{yx} = r\,.\frac{\sigma_y}{\sigma_x} \quad \Rightarrow \quad \sigma_y = \frac{b_{yx}\,.\,\sigma_x}{r} = \frac{\left(-\frac{1}{2}\right).\,2\sqrt{3}}{-\frac{\sqrt{3}}{2}} = 2$

$\therefore$ variance of $y = {\sigma_y}^2 = 4$

6. By using the data $\bar{x} = 25$, $\bar{y} = 30$, $b_{yx} = 1.6$, $b_{xy} = 0.4$, find:

(a) The regression equation of y on x.

(b) What is the most likely value of y when $x = 60$?

(c) What is the coefficient of correlation between x and y? (ISC 2017)

Solution: Given: $\quad \bar{x} = 25, \quad \bar{y} = 30, \quad b_{yx} = 1.6, \quad b_{xy} = 0.4$

(a) The regression equation of y on x is

$y - \bar{y} = b_{yx}\,(x - \bar{x})$

$\Rightarrow \quad y - 30 = 1.6\,(x - 25)$

$\Rightarrow \quad y - 30 = 1.6\,x - 40$

$\Rightarrow \quad y = 1.6x - 10$

(b) When $\quad x = 60, \quad y = 1.6 \times 60 - 10 = 86$

(c) Coefficient of correlation $r = +\sqrt{b_{xy} \times b_{yx}} = +\sqrt{1.6 \times 0.4}$

$= 0.8$

Both b_{yx} and b_{xy} are positive. So r is also positive.

7. Find the line of best fit for the following data, treating x as dependent variable (regression equation of x on y):

x	14	12	13	14	16	10	13	12
y	14	23	17	24	18	25	23	24

hence, estimate the value of x when $y = 16$. (ISC 2017)

Solution:

x	y	$x - \bar{x} = x - 13$	$y - \bar{y} = y - 21$	$(x - \bar{x})(y - \bar{y})$	$(y - \bar{y})^2 = (y - 21)^2$
14	14	1	−7	−7	49
12	23	−1	2	−2	4
13	17	0	−4	0	16
14	24	1	3	3	9
16	18	3	−3	−9	9
10	25	−3	4	−12	16
13	23	0	2	0	4
12	24	−1	3	−3	9
$\sum x = 104$	$\sum y = 168$	$\sum(x - \bar{x}) = 0$	$\sum(y - \bar{y}) = 0$	$\sum(x - \bar{x})(y - \bar{y}) = -30$	$\sum(y - \bar{y})^2 = 116$

$n = 8, \quad \bar{x} = \frac{\sum x}{n} = \frac{104}{8} = 13, \quad \bar{y} = \frac{\sum y}{n} = \frac{168}{8} = 21$

$b_{xy} = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(y - \bar{y})^2} = \frac{-30}{116} = -\frac{15}{58}$

The line of regression of x on y is given by

$$(x - \bar{x}) = b_{xy}(y - \bar{y})$$

$$\Rightarrow (x - 13) = -\frac{15}{58}(y - 21)$$

$$\Rightarrow 58x - 754 = -15y + 315$$

$$\Rightarrow 58x + 15y = 1069$$

When $y = 16$, $58x + 15 \times 16 = 1069 \Rightarrow 58x = 829$

$$\Rightarrow x = 14{\cdot}29$$

8. Find the line of regression of y on x from the following data:

x	1	2	3	4	5
y	7	6	5	4	3

Hence, estimate the value of y when $x = 6$. (ISC 2018)

Solution:

x	y	xy	x^2
1	7	7	1
2	6	12	4
3	5	15	9
4	4	16	16
5	3	15	25
$\sum x = 15$	$\sum y = 25$	$\sum xy = 65$	$\sum x^2 = 55$

$n = 5, \quad \bar{x} = \frac{\sum x}{n} = \frac{15}{5} = 3, \quad \bar{y} = \frac{\sum y}{n} = \frac{25}{5} = 5$

Regression coefficient of y on $x = b_{yx} = \frac{\sum xy - \frac{\sum x \cdot \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} = \frac{65 - \frac{15 \times 25}{5}}{55 - \frac{15^2}{5}} = \frac{-10}{10}$

$$= -1$$

The regression line of y on x is $y - \bar{y} = b_{yx}(x - \bar{x})$

$\Rightarrow \quad y - 5 = -1(x - 3)$

$\Rightarrow \quad y = -x + 8$

$\Rightarrow \quad x + y = 8$

When $x = 6, \ y = -6 + 8 = 2$

9. Find the coefficient of correlation from the following regression lines:
$x - 2y + 3 = 0, 4x - 5y + 1 = 0$ (ISC 2018)

Solution: let $x - 2y + 3 = 0$ be the regression equation of x on y and $4x - 5y + 1 = 0$ is the regression equation of y on x.

$x - 2y + 3 = 0$

$\Rightarrow \quad x = 2y - 3$

$b_{xy} = 2$

$4x - 5y + 1 = 0$

$\Rightarrow \quad 5y = 4x + 1$

$\Rightarrow \quad y = \frac{4}{5}x + \frac{1}{5}$

$\therefore \quad b_{yx} = \frac{4}{5}$

$\therefore \quad r^2 = b_{yx} \times b_{xy} = 2 \times \frac{4}{5} = \frac{8}{5} > 1$ which is not possible as $0 \leq r^2 \leq 1$. So <u>our assumption is wrong</u>.

So $x - 2y + 3 = 0$ is the regression equation of y on x and $4x - 5y + 1 = 0$ is the regression equation of x on y.

$x - 2y + 3 = 0$

$\Rightarrow \quad 2y = x + 3$

$\Rightarrow \quad y = \frac{1}{2}x + \frac{3}{2}$

$\therefore \quad b_{yx} = \frac{1}{2}$

$4x - 5y + 1 = 0$

$\Rightarrow \quad 4x = 5y - 1$

$\Rightarrow \quad x = \frac{5}{4}y - \frac{1}{4}$

$\therefore \quad b_{xy} = \frac{5}{4}$

$\therefore \quad r^2 = b_{yx} \times b_{xy} = \frac{1}{2} \times \frac{5}{4} = \frac{5}{8} \qquad \left[0 < \frac{5}{8} < 1\right]$

And $r = \sqrt{\frac{5}{8}} = 0.79$ (r is positive as both b_{yx} and b_{xy} are positive.)

10. From the given data

	x	y
Mean	6	8
standard deviation	4	6

and coefficient correlation $\frac{2}{3}$. Find:

(i) regression coefficients b_{yx} and b_{xy}.

(ii) regression line of x on y.

(iii) most likely values of x when $y = 14$. (ISC 2018)

Solution: Given

	x	y
Mean	$\bar{x} = 6$	$\bar{y} = 8$
standard deviation	$\sigma_x = 4$	$\sigma_y = 6$

$r = \frac{2}{3}$

(i) $b_{yx} = r\,\frac{\sigma_y}{\sigma_x} = \frac{2}{3} \times \frac{6}{4} = 1$ and $b_{xy} = r\,\frac{\sigma_x}{\sigma_y} = \frac{2}{3} \times \frac{4}{6} = \frac{4}{9}$

(ii) Regression line of x on y is

$$x - \bar{x} = b_{xy}(y - \bar{y})$$
$$\Rightarrow \quad x - 6 = \frac{4}{9}(y - 8)$$
$$\Rightarrow \quad x = \frac{4}{9}y - \frac{32}{9} + 6$$
$$\Rightarrow \quad x = \frac{4}{9}y + \frac{22}{9}$$

(iii) When $y = 14$, $x = \frac{4}{9} \times 14 + \frac{22}{9} = \frac{78}{9} = 8.667$

11. If $\bar{x} = 18, \bar{y} = 100, \sigma_x = 14,\ \sigma_y = 20$ and correlation coefficient $r_{xy} = 0.8$, find the regression equation of y on x. (ISC 2019)

Solution: Given: $\bar{x} = 18,\ \bar{y} = 100,\ \sigma_x = 14,\ \sigma_y = 20$ and $r_{xy} = 0.8$

$$b_{yx} = r\,\frac{\sigma_y}{\sigma_x} = 0.8 \times \frac{20}{14} = \frac{8}{7}$$

$\therefore$ Regression equation of y on x is

$$y - \bar{y} = b_{yx}(x - \bar{x})$$
$$\Rightarrow \quad y - 100 = \frac{8}{7}(x - 18)$$
$$\Rightarrow \quad 7y - 700 = 8x - 144$$
$$\Rightarrow \quad 8x - 7y + 556 = 0$$

12. The following results were obtained with respect to two variables x and y:
$\sum x = 15,\ \sum y = 25, \sum xy = 83,\ \sum x^2 = 55, \sum y^2 = 135$ and $n = 5$

(i) Find the regression coefficient b_{xy}

(ii) Find the regression equation of x on y. (ISC 2019)

Solution: Given: $\sum x = 15,\ \sum y = 25, \sum xy = 83,\ \sum x^2 = 55, \sum y^2 = 135$ and $n = 5$

(i) Regression coefficient of x on $y = b_{xy} = \dfrac{\sum xy - \frac{\sum x \,.\, \sum y}{n}}{\sum y^2 - \frac{(y)^2}{n}}$

$$= \frac{83 - \frac{15 \times 25}{5}}{135 - \frac{25^2}{5}} = \frac{8}{10} = \frac{4}{5}$$

(ii) The regression equation of x on y is

$$x - \bar{x} = b_{xy}(y - \bar{y})$$
$$\Rightarrow \quad x - 3 = \frac{4}{5}(y - 5)$$
$$\Rightarrow \quad 5x - 15 = 4y - 20$$
$$\Rightarrow \quad 5x - 4y + 5 = 0$$

13. Find the equation of the regression line of y on x, if the observations (x, y) are as follows:

(1,4), (2,8), (3,2), (4,12), (5,10), (6,14), (7,16), (8,6), (9,18).

Also, find the estimated value of y when $x = 14$. (ISC 2019)

Solution:

x	y	xy	x^2
1	4	4	1
2	8	16	4
3	2	6	9
4	12	48	16
5	10	50	25
6	14	84	36
7	16	112	49
8	6	48	64
9	18	162	81
$\sum x = 45$	$\sum y = 90$	$\sum xy = 530$	$\sum x^2 = 285$

$\bar{x} = \frac{\sum x}{n} = \frac{45}{9} = 5$ and $\bar{y} = \frac{\sum y}{n} = \frac{90}{n} = 10$

Regression coefficient of y on $x = b_{yx} = \frac{\sum xy - \frac{\sum x \cdot \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$

$= \frac{530 - \frac{45 \times 90}{9}}{285 - \frac{45^2}{9}} = \frac{80}{60} = \frac{4}{3}$

$\therefore$ Regression equation of y on x is

$y - \bar{y} = b_{yx}(x - \bar{x})$

$\Rightarrow \quad y - 10 = \frac{4}{3}(x - 5)$

$\Rightarrow \quad 3y - 30 = 4x - 20$

$\Rightarrow \quad 4x - 3y + 10 = 0$

When $x = 14$, $4 \times 14 - 3y + 10 = 0 \quad \Rightarrow \quad y = 22$

14. For the lines of regression $4x - 2y = 4$ and $2x - 3y + 6 = 0$, find the mean of $'x'$, mean of $'y'$. (ISC 2020)

Solution:

Given $4x - 2y = 4$ --------- (1)

And $2x - 3y + 6 = 0$ ---------(2)

Multiplying equation (2) by 2, $4x - 6y + 12 = 0$

Equation (1) is $4x - 2y - 4 = 0$

$- \quad + \quad +$

$-4y + 16 = 0$

$\Rightarrow \quad y = 4$

Putting value of $y = 4$ in (1), $4x - 2 \times 4 = 4 \quad \Rightarrow x = 3$

$\therefore \quad \bar{x} = 3$ and $\bar{y} = 4$

15. The correlation coefficient between x and y Is 0.6. If the variance of x is 225, the variance of y is 400, mean of x is 10 and mean of y is 20, find:

(i) The equations of 2 regression lines

(ii) The expected value of y when $x = 2$. (ISC 2020)

Solution: Given

	x	y
Mean	$\bar{x} = 10$	$\bar{y} = 20$
variance	$\sigma_x^2 = 2\ 25$	$\sigma_y^2 = 400$

Coefficient of correlation between x and y = r = 0.6

$\therefore$ Standard deviation:

$$\sigma_x = \sqrt{225} = 15$$

$$\sigma_y = \sqrt{400} = 20$$

Now $b_{xy} = r\ \frac{\sigma_x}{\sigma_y} = 0.6 \times \frac{15}{20} = \frac{9}{20}$

And $b_{yx} = r\ \frac{\sigma_y}{\sigma_x} = 0.6 \times \frac{20}{15} = \frac{4}{5}$

(i) The regression equation of x on y is

$$x - \bar{x} = b_{xy}\,(y - \bar{y})$$

$$\Rightarrow \quad x - 10 = \frac{9}{20}\,(y - 20)$$

$$\Rightarrow \quad 20\,x - 200 = 9y - 180$$

$$\Rightarrow \quad 20\,x - 9y = 20$$

The regression equation of y on x is

$$y - \bar{y} = b_{yx}\,(x - \bar{x})$$

$$\Rightarrow \quad y - 20 = \frac{4}{5}(x - 10)$$

$$\Rightarrow \quad 5y - 100 = 4x - 40$$

$$\Rightarrow \quad 4x - 5y + 60 = 0$$

(ii) The regression equation of y on x is

$$4x - 5y + 60 = 0$$

When $x = 2$, $\quad 8 - 5y + 60 = 0 \ \Rightarrow \ 5y = 68 \quad \Rightarrow y = \frac{68}{5} = 13.6$

16. Find the regression coefficients b_{yx} and b_{xy} and correlation coefficient $'r'$ for the following data:

(2,8), (6, 8), (4,5), (7,6), (5,2) (ISC 2020)

Solution:

x	y	xy	x^2	y^2
2	8	16	4	64
6	8	48	36	64
4	5	20	16	25
7	6	42	49	36
5	2	10	25	4
$\sum x = 24$	$\sum y = 29$	$\sum xy = 136$	$\sum x^2 = 130$	$\sum y^2 = 193$

Regression coefficient of y on $x = b_{yx} = \frac{\sum x\,y - \frac{\sum x\,.\,\sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} = \frac{136 - \frac{24 \times 29}{5}}{130 - \frac{24^2}{5}} = \frac{-3.2}{14.8}$

$$= -\frac{8}{37} = -0.216$$

Regression coefficient of x on $y = b_{xy} = \frac{\sum x\,y - \frac{\sum x\,.\,\sum y}{n}}{\sum y^2 - \frac{(\sum y)^2}{n}} = \frac{136 - \frac{24 \times 29}{5}}{193 - \frac{29^2}{5}} = \frac{-3.2}{24.8}$

$$= -\frac{4}{31} = -0.129$$

Correlation coefficient $r = -\sqrt{b_{yx} \times b_{xy}}$

$$= -\sqrt{(-0.216)(-0.129)} = -0.16669$$

17. Choose the correct option: (ISC SEM-2 2022)

If the two regression coefficients b_{xy} and b_{yx} are -0.8 and -0.2 respectively, then the correlation coefficient (ρ) will be,

(a) 0.16
(b) – 0.16
(c) 0.4
(d) – 0.4

Solution: **(d)**

$\rho^2 = b_{xy} \times b_{yx} = (-0.8) \times (-0.2) = 0.16$

$\therefore \quad \rho = -\sqrt{0.16} = -0.4$

ρ is negative as both b_{xy} and b_{yx} are negative.

18. Choose the correct option: (ISC SEM-2 2022)

The line of regression of y on x is, $4x - 5y + 33 = 0$ and the line of regression of x on y is, $20x - 9y - 107 = 0$, then the value of x when $y = 7$ is,

(a) 8.5
(b) – 8.5
(c) 0.5
(d) – 0.5

Solution: **(a)**

The line of regression of x on y is, $20x - 9y - 107 = 0$

$\therefore \quad 20x = 9y + 107$

When $y = 7$, $\quad 20x = 9 \times 7 + 107 = 170$

$\therefore \quad x = \frac{170}{20} = \frac{17}{2} = 8.5$

19. The mean and standard deviation of the two variables x and y are given as $\bar{x} = 6,\ \bar{y} = 8,\ \sigma_x = 4,\ \sigma_y = 12$. The correlation coefficient is given as $r = \frac{2}{3}$.

Find the regression line of x on y. (ISC SEM-2 2022)

Solution:

We have to find out the regression line of x on y.

$b_{xy} = r\,\frac{\sigma_x}{\sigma_y} = \frac{2}{3} \times \frac{4}{12} = \frac{2}{9}$

The regression line of x on y is

$x - \bar{x} = b_{xy}(y - \bar{y})$

$\Rightarrow \quad x - 6 = \frac{2}{9}(y - 8)$

$\Rightarrow \quad 9x - 54 = 2y - 16$

$\Rightarrow \quad 9x - 2y = 38$